

**COMMUNICATION SUPPORT SYSTEM AND PHOTOGRAPHING EQUIPMENT**

Patent Number: JP2002261966  
Publication date: 2002-09-13  
Inventor(s): HOSHINO HIROSHI; SUZUKI TAKEYUKI; NISHIYAMA TAKASHI  
Applicant(s): MATSUSHITA ELECTRIC WORKS LTD  
Requested Patent: ☐ JP2002261966  
Application Number: JP20010222210 20010723  
Priority Number(s):  
IPC Classification: H04M11/00; G08B21/02; G08B25/08; G10L13/00; G10L15/00; G10L15/22; H04N7/15; H04N7/18  
EC Classification:  
Equivalents:

---

**Abstract**

---

**PROBLEM TO BE SOLVED:** To enable to provide a mental care for loneliness by simulating human conversations and to watch whether there is anything wrong by transmitting an image to a mobile terminal within the extent not invading the privacy.

**SOLUTION:** A robot 1 comprises a conversation processing means 21 to which a microphone 11 and a speaker 12 are connected and which enables voice conversations with the other party using a natural language, an image processing means 22 which is connected by a TV camera 13 for photographing the other party and outputs an image data which is a digitized image photographed by the TV camera 13, and a communication means 23 for sending out the image data onto a network. The image data is transferred to an image server 4 via a communication line 3. By accessing the image server 4 from a mobile terminal 2 which is allowed to access the image server, the image data can be displayed on the mobile terminal 2.

---

Data supplied from the esp@cenet database - I2

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2002-261966  
(P2002-261966A)

(43) 公開日 平成14年9月13日 (2002.9.13)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テ-マ-ト*(参考)
H 0 4 M 11/00	3 0 2	H 0 4 M 11/00	5 C 0 5 4
G 0 8 B 21/02		G 0 8 B 21/02	5 C 0 6 4
25/08		25/08	E 5 C 0 8 6
G 1 0 L 13/00		H 0 4 N 7/15	5 C 0 8 7
15/00		7/18	D 5 D 0 1 5

審査請求 未請求 請求項の数27 O L (全 22 頁) 最終頁に続く

(21) 出願番号 特願2001-222210(P2001-222210)  
(22) 出願日 平成13年7月23日 (2001.7.23)  
(31) 優先権主張番号 特願2000-274120(P2000-274120)  
(32) 優先日 平成12年9月8日 (2000.9.8)  
(33) 優先権主張国 日本 (J P)  
(31) 優先権主張番号 特願2000-393678(P2000-393678)  
(32) 優先日 平成12年12月25日 (2000.12.25)  
(33) 優先権主張国 日本 (J P)

(71) 出願人 000005832  
松下電工株式会社  
大阪府門真市大字門真1048番地  
(72) 発明者 星野 洋  
大阪府門真市大字門真1048番地松下電工株式会社内  
(72) 発明者 鈴木 健之  
大阪府門真市大字門真1048番地松下電工株式会社内  
(74) 代理人 100087767  
弁理士 西川 恵清 (外1名)

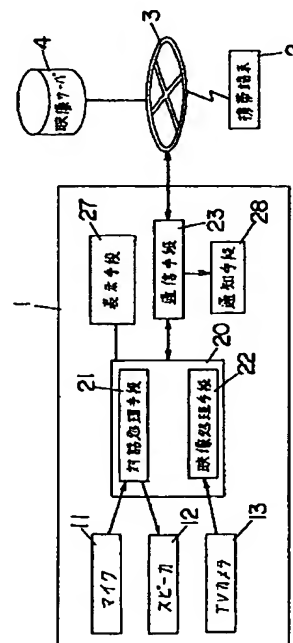
最終頁に続く

(54) 【発明の名称】 コミュニケーション支援システムおよび撮影装置

(57) 【要約】

【課題】人との対話を模擬することによって人の孤独感の精神的なケアを可能とするとともに、プライバシーを侵害しない範囲で対話者の画像を携帯端末に転送して対話者の異常の有無を監視可能とする。

【解決手段】ロボット1は、マイクロホン11とスピーカ12とが接続され自然言語を用いた音声による対話を対話者との間で行う対話処理手段21と、対話者を撮像するTVカメラ13が接続されTVカメラ13で撮像された映像をデジタル化した映像データを出力する映像処理手段22と、映像データをネットワークに送出する通信手段23とを備える。映像データは通信回線3を通して映像サーバ4に転送され、閲覧が許可された携帯端末2から映像サーバ4をアクセスすることによって、映像データを携帯端末2に表示することができる。



## 【特許請求の範囲】

【請求項1】 音声入力装置と音声出力装置とが接続され自然言語を用いた音声による対話を対話者との間で行う対話処理手段と、対話者を撮像する映像入力装置が接続され映像入力装置で撮像された映像をデジタル化した映像データを出力する映像処理手段と、映像データをネットワークに送出する通信手段と、ネットワークを通して対話者の映像データの閲覧が許可されている携帯端末と、映像データを携帯端末に表示可能なフォーマットに変換する変換手段とを備えることを特徴とするコミュニケーション支援システム。

【請求項2】 前記変換手段がネットワークを介して転送された映像データを蓄積する映像サーバに設けられ、映像サーバでは携帯端末の所持者に許容される映像データのみを選択して携帯端末に転送することを特徴とする請求項1記載のコミュニケーション支援システム。

【請求項3】 前記変換手段が映像処理手段と通信手段との一方に設けられていることを特徴とする請求項1記載のコミュニケーション支援システム。

【請求項4】 前記映像処理手段が前記携帯端末からの指示により対話者の映像データを通信手段を通してネットワークに送出させることを特徴とする請求項1ないし請求項3のいずれか1項に記載のコミュニケーション支援システム。

【請求項5】 前記映像処理手段が前記対話処理手段を通して対話者が指示した命令にตอบสนองして動作することを特徴とする請求項1ないし請求項4のいずれか1項に記載のコミュニケーション支援システム。

【請求項6】 前記映像処理手段が対話者からの指示により動作しているか前記携帯端末からの指示によって動作しているかを被指示側に報知する報知手段を備えることを特徴とする請求項1ないし請求項5のいずれか1項に記載のコミュニケーション支援システム。

【請求項7】 対話者の接触を検出する接触センサが付加され、接触センサにより対話者の接触が検出されると前記対話処理手段が起動することを特徴とする請求項1ないし請求項6のいずれか1項に記載のコミュニケーション支援システム。

【請求項8】 前記映像入力手段によって対話者の映像が撮影されていることを対話者に示す手段を備えることを特徴とする請求項1ないし請求項7のいずれか1項に記載のコミュニケーション支援システム。

【請求項9】 前記対話処理手段が、対話者による音声入力を確認するように対話者に質問する機能を有し、対話者による肯定の応答を受けると対話者による音声入力を確定することを特徴とする請求項1ないし請求項8のいずれか1項に記載のコミュニケーション支援システム。

【請求項10】 前記対話処理手段が、対話者による肯定の応答を教師信号として対話者の発生する音声学習

することを特徴とする請求項9記載のコミュニケーション支援システム。

【請求項11】 前記携帯端末で閲覧した映像データを前記ネットワーク上の他の端末に転送する手段を設けたことを特徴とする請求項1ないし請求項10のいずれか1項に記載のコミュニケーション支援システム。

【請求項12】 前記携帯端末から映像データの閲覧が要求されたことを前記対話者に通知する通知手段を備えることを特徴とする請求項1ないし請求項11のいずれか1項に記載のコミュニケーション支援システム。

【請求項13】 前記通知手段が前記携帯端末の識別情報を用いて映像データの閲覧者を通知する機能を備えることを特徴とする請求項12記載のコミュニケーション支援システム。

【請求項14】 前記通知手段が映像データの閲覧者ごとに割り当てた報知部を備えることを特徴とする請求項13記載のコミュニケーション支援システム。

【請求項15】 音声入力装置と音声出力装置とが接続され自然言語を用いた音声による対話を対話者との間で行う対話処理手段と、対話者を撮像する映像入力装置が接続され映像入力装置で撮像された映像をデジタル化した映像データを出力する映像処理手段と、映像データをネットワークに送出する通信手段と、ネットワークを通して対話者の映像データの閲覧が許可されている携帯端末に表示可能なフォーマットとなるように映像データを変換する変換手段とを備えることを特徴とする撮影装置。

【請求項16】 現在日時を計時する時計手段と、対話者が撮像されたときに時計手段で計時されている日時を文字として映像データに合成して通信手段から送出させるスーパーインポーズ処理手段とが付加されたことを特徴とする請求項15記載の撮影装置。

【請求項17】 前記通信手段を通してネットワークに送出される映像データおよび通信の成功・失敗の結果を前記時計手段で計時されている日時に対応付けて順次記録して蓄積するとともに無給電で記録内容を保持する記憶手段が付加されていることを特徴とする請求項16記載の撮影装置。

【請求項18】 前記映像入力装置の傾きを検出する姿勢センサが付加され、前記映像処理手段では映像入力装置により撮像された映像を正立させるように姿勢センサにより検出された傾きに基づいて映像を補正することを特徴とする請求項15ないし請求項17のいずれか1項に記載の撮影装置。

【請求項19】 前記映像入力装置の視野の方向を変更可能とする首振り駆動手段と、前記映像入力装置の中央付近に対話者を撮像するように規定したテンプレートが登録可能であって、映像入力装置で撮像された映像とテンプレートとを照合して類似度が大きくなるように首振り駆動手段を制御するパターン処理手段とが付加されて

いることを特徴とする請求項15ないし請求項18のいずれか1項に記載の撮影装置。

【請求項20】 前記映像入力装置の視野の方向を指向性を有する方向に一致させた第2の音声入力装置と、前記映像入力装置の視野の方向と第2の音声入力装置が指向性を有する方向とを同時に変更可能とする首振り駆動手段と、前記第2の音声入力装置の出力の音レベルが大きくなる向きに首振り駆動手段を制御する音レベル処理手段とが付加されていることを特徴とする請求項15ないし請求項18のいずれか1項に記載の撮影装置。

【請求項21】 前記映像入力装置の視野の方向に存在する音源に対して出力の位相差がゼロになるように配置した複数の第3の音声入力装置と、第3の音声入力装置の位相差を小さくする向きに首振り駆動手段を制御する位相差検出手段とを付加したことを特徴とする請求項15ないし請求項18のいずれか1項に記載の撮影装置。

【請求項22】 前記映像入力装置の前方に前記映像入力装置の光軸方向に直交させた形で配置したハーフミラーが付加されていることを特徴とする請求項15ないし請求項21のいずれか1項に記載の撮影装置。

【請求項23】 外部のディスプレイ装置が接続可能であって前記映像入力装置により撮像された映像データをディスプレイ装置に表示可能とする映像出力手段が付加されていることを特徴とする請求項15ないし請求項21のいずれか1項に記載の撮影装置。

【請求項24】 前記対話者が抱えることを可能とする外殻を有し、前記対話者が抱えたときに対話者の心拍音を検出可能な心拍センサと、心拍センサにより検出された心拍音を記録する心拍記録装置とが付加されていることを特徴とする請求項15ないし請求項23のいずれか1項に記載の撮影装置。

【請求項25】 前記携帯端末から映像データの閲覧が要求されたことを前記対話者に通知する通知手段を備えることを特徴とする請求項15ないし請求項24のいずれか1項に記載の撮影装置。

【請求項26】 前記通知手段が前記携帯端末の識別情報を用いて映像データの閲覧者を通知する機能を備えることを特徴とする請求項25記載の撮影装置。

【請求項27】 前記通知手段が映像データの閲覧者ごとに割り当てた報知部を備えることを特徴とする請求項26記載の撮影装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、コミュニケーション支援システムおよび撮影装置に関するものである。

【0002】

【従来の技術】一般に、高齢者は外出の機会が少なく他者との会話が少ないから、家族が会話の相手になっていることが多い。また、高齢者では体力の衰えから家庭内

での事故も多く、事故を未然に防止したり事故があれば迅速に対処するために、家族が常に見守っていることも必要である。このような事情から、高齢者とともに生活している家族は高齢者を家に残して外出することが難しく、旅行など言うに及ばず、買い物や習い事すらままならないことが多い。

【0003】一方、最近では、あたかも意思を持つかのように行動するペットロボットと称する種類のロボットが製品化されており、この種のロボットとしては、たとえば犬型のペットロボットが「AIBO」（ソニー社の商品名）という名称で製品化されている。このロボットは、人の声の調子や表情や触れ方などを監視して学習することによって、従来の機械装置のように刺激に対する応答が定式化せず刺激に対する応答が状況に応じて変化することになり、成長しているかのように振る舞ったり、人とコミュニケーションをしているかのように錯覚させることが可能になり、一種の癒し機能を有するものになっている。ただし、このペットロボットは、犬の動きを模倣してエンターテインメントを指向するものであり、高齢者とのコミュニケーションの支援を目的とするものではない。

【0004】一方、高齢者とのコミュニケーションを主たる目的としたペットロボットとしては、松下電器産業株式会社から猫型あるいは熊型のロボットが提案されており、このペットロボットは高齢者の快適な生活に貢献するために、簡単な日常会話とペットのような振る舞いが行えるほか、遠隔地からペットロボットの使用状況を間接的に把握する機能も備えている。したがって、このペットロボットは高齢者の話し相手となって高齢者の精神的なケアを行い、また高齢者の独居生活の安全の確認を遠隔地から行うことが可能になっている。このペットロボットは、通信回線（双方向CATV施設にデジタル通信システム技術およびモバイル通信技術を複合した通信回線）を介して福祉サービス支援センタに設置されたセンタ設備と接続されるものであって、独居している高齢者の安全を確認する際には、ペットロボットに発話させるメッセージをセンタ設備から通信回線を通してペットロボットに伝送し、ペットロボットから高齢者に向かってメッセージを音声によって伝える。このときの高齢者の応答をセンタ設備に返送し、センタ設備に設けた推論装置によって生活の状況を推論することによって、高齢者の安全の確認を行う。推論装置では、たとえばペットロボットの発話に対して高齢者の応答がないときに、異常が生じている可能性があるなどと推論するのである。

【0005】

【発明が解決しようとする課題】上述したペットロボットには、生き物のように反応する機能があり、前者のペットロボットには対話機能がないものの、後者のペットロボットは対話機能を備えているから、高齢者の孤独感

を軽減することが可能と考えられる。しかしながら、後者のペットロボットは、公共施設である福祉サービス支援センタに通信回線を介して接続されるものであるから、高齢者の生活のプライバシーが侵害されることのないように、高齢者の生活の様子を撮影した映像を転送する機能は備えていない。したがって、異常が生じていると推論したときにも、現場に赴かなければどのような異常が生じているかを確認することができず、病気やけがのように迅速な対処が要求される事態に対しての処置に遅れが生じる可能性がある。

【0006】家族が外出したり旅行したりする際に家に残した高齢者の様子を確認するのであれば、テレビ電話に代表されるように画像の転送が可能な機器を用いることが考えられるが、この種の機器は通話時において映像の転送も可能としているに過ぎず、通話時以外に高齢者の会話の相手になる機能はなく、通話時以外に高齢者の精神的なケアを行う機能はない。このような事情は高齢者を抱える家庭だけではなく、病人を抱える家庭においても同様である。

【0007】本発明は上記事由に鑑みて為されたものであり、その目的は、人との対話を模擬することによって人の孤独感の精神的なケアを可能とするとともに、プライバシーを侵害しない範囲で対話者の画像を携帯端末に転送して対話者の異常の有無を監視可能としたコミュニケーション支援システムおよび撮影装置を提供することにある。

【0008】

【課題を解決するための手段】請求項1の発明は、音声入力装置と音声出力装置とが接続され自然言語を用いた音声による対話を対話者との間で行う対話処理手段と、対話者を撮像する映像入力装置が接続され映像入力装置で撮像された映像をデジタル化した映像データを出力する映像処理手段と、映像データをネットワークに送出する通信手段と、ネットワークを通して対話者の映像データの閲覧が許可されている携帯端末と、映像データを携帯端末に表示可能なフォーマットに変換する変換手段とを備えるものである。この構成によれば、対話処理手段を用いて人との対話を模擬するから対話者は孤独感に対する精神的なケアがなされる。また、対話者の映像データの閲覧は許可されている携帯端末のみで可能であるから、携帯端末を対話者の家族のみが携帯できるようにするように使用することによって、他人のプライバシーを侵害しない範囲で対話者の画像を携帯端末に転送して対話者の異常の有無を監視可能とすることができる。その結果、高齢者や病人を抱える家庭においても高齢者や病人の様子を携帯端末により遠方から監視することができ、安心して外出することができるようになる。

【0009】請求項2の発明は、請求項1の発明において、前記変換手段がネットワークを介して転送された映像データを蓄積する映像サーバに設けられ、映像サーバ

では携帯端末の所持者に許容される映像データのみを選択して携帯端末に転送するものである。この構成によれば、対話者の映像データを蓄積する映像サーバを設けたことによって、過去に遡って映像データを閲覧することが可能になる。

【0010】請求項3の発明は、請求項1の発明において、前記変換手段が映像処理手段と通信手段との一方に設けられているものである。この構成によれば、映像サーバを設けることなく対話者の映像データを携帯端末から閲覧することができ、システム構成が簡単である。

【0011】請求項4の発明は、請求項1ないし請求項3の発明において、前記映像処理手段が前記携帯端末からの指示により対話者の映像データを通信手段を通してネットワークに送出させるものである。この構成によれば、携帯端末から映像データの送出を指示するから、携帯端末の所持者が要求したときの映像を見ることができ

る。【0012】請求項5の発明は、請求項1ないし請求項4の発明において、前記映像処理手段が前記対話処理手段を通して対話者が指示した命令に回答して動作するものである。この構成によれば、対話者が自身の意思で映像を撮影させるから、たとえば対話者が異常を感じたときに映像データを送って携帯端末の所持者に知らせることが可能になる。

【0013】請求項6の発明は、請求項1ないし請求項5の発明において、前記映像処理手段が対話者からの指示により動作しているか前記携帯端末からの指示によって動作しているかを被指示側に報知する報知手段を備えるものである。この構成によれば、対話者からと携帯端末からとのどちらからも映像処理手段の制御を可能としながらも、両者の指示が混乱することがない。

【0014】請求項7の発明は、請求項1ないし請求項6の発明において、対話者の接触を検出する接触センサが付加され、接触センサにより対話者の接触が検出されると前記対話処理手段が起動するものである。この構成によれば、対話者が対話しようとするときに接触センサに接触することによって対話処理手段が起動されるから、周囲の雑音や対話者以外の他人の音声によって対話処理手段が誤動作するのを防止することができる。

【0015】請求項8の発明は、請求項1ないし請求項7の発明において、前記映像入力手段によって対話者の映像が撮影されていることを対話者に示す手段を備えるものである。この構成によれば、対話者がいつ撮影されているかを知ることができるから、プライバシーの都合で撮影されたくないときには撮影されない場所に移動するなどの対処が可能になる。

【0016】請求項9の発明は、請求項1ないし請求項8の発明において、前記対話処理手段が、対話者による音声入力を確認するように対話者に質問する機能を有し、対話者による肯定の応答を受けると対話者による音

声入力を確定するものである。この構成によれば、復唱によって対話者の音声入力を確認するから、対話者の発生した音声を実際に理解することが可能になる。

【0017】請求項10の発明は、請求項9の発明において、前記対話処理手段が、対話者による肯定の応答を教師信号として対話者の発生する音声进行学习するものである。この構成によれば、対話者との対話によって対話者の発生する音声进行学习することで認識率が高くなる。

【0018】請求項11の発明は、請求項1ないし請求項10の発明において、前記携帯端末で閲覧した映像データを前記ネットワーク上の他の端末に転送する手段を設けたものである。この構成によれば、映像データを他の端末に転送することにより、たとえば携帯端末の所持者が対話者の異常に気づいたときに映像データを医者に送ることで対処のアドバイスを受けるなどの使用が可能になる。

【0019】請求項12の発明は、請求項1ないし請求項11の発明において、前記携帯端末から映像データの閲覧が要求されたことを前記対話者に通知する通知手段を備えるものである。この構成によれば、対話者である高齢者や病人を抱える家庭においても家族が携帯端末を用いて外出先から高齢者や病人の様子を監視したときに、高齢者や病人に対しては通知手段によって家族が確認したことが通知されるから、高齢者や病人にとっては家族が確認してくれたことを知ることができ安心感が得られる。

【0020】請求項13の発明は、請求項12の発明において、前記通知手段が前記携帯端末の識別情報を用いて映像データの閲覧者を通知する機能を備えるものである。この構成によれば、映像データを誰が確認したかを知ることができるから、より安心することができる。

【0021】請求項14の発明は、請求項13の発明において、前記通知手段が映像データの閲覧者ごとに割り当てた報知部を備えるものである。この構成によれば、文字データなどによらずに報知部を視認すれば閲覧者を識別することができるから、閲覧者が映像データを確認したことを知るだけでなく誰が閲覧したかも直感的に知ることができる。

【0022】請求項15の発明は、音声入力装置と音声出力装置とが接続され自然言語を用いた音声による対話を対話者との間で行う対話処理手段と、対話者を撮像する映像入力装置が接続され映像入力装置で撮像された映像をデジタル化した映像データを出力する映像処理手段と、映像データをネットワークに送出する通信手段と、ネットワークを通して対話者の映像データの閲覧が許可されている携帯端末に表示可能なフォーマットとなるように映像データを変換する変換手段とを備えるものである。この構成によれば、対話処理手段を用いて人との対話を模擬するから対話者は孤独感に対する精神的なケアがなされる。また、対話者の映像データの閲覧は許可さ

れている携帯端末のみで可能であるから、携帯端末を対話者の家族のみが携帯できるようにするように使用することによって、他人のプライバシーを侵害しない範囲で対話者の画像を携帯端末に転送して対話者の異常の有無を監視可能とすることができる。その結果、高齢者や病人を抱える家庭においても高齢者や病人の様子を携帯端末により遠方から監視することができ、安心して外出することができるようになる。

【0023】請求項16の発明は、請求項15の発明において、現在日時を計時する時計手段と、対話者が撮像されたときに時計手段で計時されている日時を文字として映像データに合成して通信手段から送出させるスーパーインポーズ処理手段とが付加されたものである。この構成によれば、各映像データに映像の撮像された日時が添付されているから、映像が撮像された日時と映像データを携帯端末によって閲覧する日時とに大きな時間差があるようなときに、閲覧者は映像データに添付されている日時を確認することによって、最新の映像を求めるなどの対処が可能になる。

【0024】請求項17の発明は、請求項16の発明において、前記通信手段を通してネットワークに送出される映像データおよび通信の成功・失敗の結果を前記時計手段で計時されている日時に対応付けて順次記録して蓄積するとともに無給電で記録内容を保持する記憶手段が付加されているものである。この構成によれば、通信手段から映像データが送出されず携帯端末によって映像データを確認することができないような場合にも、記憶手段に蓄積された映像データによって対話者の健康状態などを後日確認することができ、また通信の成功・失敗の結果を送信履歴として記憶手段に格納しているから、メンテナンス時に通信手段の異常の有無などを容易に知ることができる。

【0025】請求項18の発明は、請求項15ないし請求項17の発明において、前記映像入力装置の傾きを検出する姿勢センサが付加され、前記映像処理手段では映像入力装置により撮像された映像を正立させるように姿勢センサにより検出された傾きに基づいて映像を補正するものである。この構成によれば、姿勢センサにより得られている情報に基づいて映像の傾きを除去するように補正するから、映像入力装置が傾いているような場合、たとえば水平ではない場所に撮影装置が配置されている場合や対話者が撮影装置を持ち上げているような場合においても、撮影装置が水平面上に配置されている状態と同様の映像を得ることができる。

【0026】請求項19の発明は、請求項15ないし請求項18の発明において、前記映像入力装置の視野の方向を変更可能とする首振り駆動手段と、前記映像入力装置の中央付近に対話者を撮像するように規定したテンプレートが登録可能であって、映像入力装置で撮像された映像とテンプレートとを照合して類似度が大きくなるよ



うに首振り駆動手段を制御するパターン処理手段とが付加されているものである。この構成によれば、映像入力装置により撮像された映像とテンプレートとの類似度が大きくなるように映像入力装置の視野の向きが制御されるから、適切なテンプレートが設定されていれば対話者を映像入力装置の視野の中央付近に捕捉するように映像入力装置の向きを調節することが可能になる。

【0027】請求項20の発明は、請求項15ないし請求項18の発明において、前記映像入力装置の視野の方向を指向性を有する方向に一致させた第2の音声入力装置と、前記映像入力装置の視野の方向と第2の音声入力装置が指向性を有する方向とを同時に変更可能とする首振り駆動手段と、前記第2の音声入力装置の出力の音レベルが大きくなる向きに首振り駆動手段を制御する音レベル処理手段とが付加されているものである。この構成によれば、音レベルによって対話者の存在する方向を推定することができるから、映像入力装置の視野を音レベルの大きくなる向きに調節することによって、映像入力装置の視野内に対話者を捕捉できる可能性が高くなる。

【0028】請求項21の発明は、請求項15ないし請求項18の発明において、前記映像入力装置の視野の方向に存在する音源に対して出力の位相差がゼロになるように配置した複数の第3の音声入力装置と、第3の音声入力装置の位相差を小さくする向きに首振り駆動手段を制御する位相差検出手段とを付加したものである。この構成によれば、複数の音声入力装置の位相差を小さくするように映像入力装置の視野を調節することによって、映像入力装置の視野内に対話者を捕捉することが可能になる。

【0029】請求項22の発明は、請求項15ないし請求項21の発明において、前記映像入力装置の前方に前記映像入力装置の光軸方向に直交させた形で配置したハーフミラーが付加されているものである。この構成によれば、対話者がハーフミラーに自身の姿を映した状態では映像入力装置の正面に対話者が存在することになるから、対話者の意思によって対話者を映像入力装置で確実に撮像することが可能になる。

【0030】請求項23の発明は、請求項15ないし請求項21の発明において、外部のディスプレイ装置が接続可能であって前記映像入力装置により撮像された映像データをディスプレイ装置に表示可能とする映像出力手段が付加されているものである。この構成によれば、対話者は、映像入力装置により撮像されている映像を映像出力手段に接続したディスプレイ装置によってその場で確認できるから、対話者は自身の映り具合を確認することができるだけでなく周囲の様子も確認して適切な映像が撮像されているか否かを確認することができる。

【0031】請求項24の発明は、請求項15ないし請求項23の発明において、前記対話者が抱えることを可能とする外殻を有し、前記対話者が抱えたときに対話者

の心拍音を検出可能な心拍センサと、心拍センサにより検出された心拍音を記録する心拍記録装置とが付加されているものである。この構成によれば、映像によって対話者の健康状態を知るだけではなく、心拍音によっても対話者の健康状態を確認することが可能になり、対話者に関して得られる情報を複合することによって対話者の健康状態をより正確に知ることが可能になる。しかも、心拍センサとして対話者が抱えたときに心拍音を集音する構成を採用しているから、手間をかけることなく心拍音の集音が可能になる。

【0032】請求項25の発明は、請求項15ないし請求項24の発明において、前記携帯端末から映像データの閲覧が要求されたことを前記対話者に通知する通知手段を備えるものである。この構成によれば、対話者である高齢者や病人を抱える家庭においても家族が携帯端末を用いて外出先から高齢者や病人の様子を監視したときに、高齢者や病人に対しては通知手段によって家族が確認したことが通知されるから、高齢者や病人にとっては家族が確認してくれたことを知ることができ安心感が得られる。

【0033】請求項26の発明は、請求項25の発明において、前記通知手段が前記携帯端末の識別情報を用いて映像データの閲覧者を通知する機能を備えるものである。この構成によれば、映像データを誰が確認したかを知ることができるから、より安心することができる。

【0034】請求項27の発明は、請求項26の発明において、前記通知手段が映像データの閲覧者ごとに割り当てた報知部を備えるものである。この構成によれば、文字データなどによらずに報知部を視認すれば閲覧者を識別することができるから、閲覧者が映像データを確認したことを知るだけでなく誰が閲覧したかも直感的に知ることができる。

【0035】

【発明の実施の形態】（第1の実施の形態）まず、本発明に用いるシステムの全体構成について説明する。本実施形態は、図1に示すように、高齢者や病人（以下では、ロボット1と対話する者という意味で単に「対話者」と呼ぶ）がいる宅内に配置されるロボット1と、このロボット1にインターネットのようなネットワークを含む通信回線3を介して接続される携帯端末2とを用いる。ロボット1は、音声による対話の機能と、周囲を撮像する機能と、映像を含むデータの通信機能とを備えている。本実施形態においては、ロボット1が大型の外観を呈するものとして説明するが、動物などの動きを模擬することは必須要件ではなく、上述した機能を備えているものであれば、外観形状や動きについてはとくに問題にする必要はない。本実施形態では、携帯端末2として画像表示機能を有する携帯電話を想定する。また、通信回線3には、ロボット1により撮像した映像を一時的に蓄積する映像サーバ4が接続される。

【0036】ロボット1は、音声による対話を行うために、耳としてのマイクロホン11（音声入力装置）と口としてのスピーカ12（音声出力装置）とを備え、またロボット1の周囲の撮像のために、目としての小型のTVカメラ13（映像入力装置）を備える。マイクロホン11およびスピーカ12は対話処理手段21に接続されており、対話処理手段21ではマイクロホン11およびスピーカ12を通して自然言語による対話処理が可能になっている。一方、TVカメラ13は映像処理手段22に接続され、映像処理手段22ではTVカメラ13により撮像された映像の輝度やコントラストを調整したり、映像信号をデジタル化したりする。対話処理手段21および映像処理手段22は、マイコンを主構成要素とする信号処理部20を構成する。

【0037】映像処理手段22の後段には通信手段23が接続され、通信手段23では映像信号を通信回線3に送出することができるデータフォーマットに圧縮して通信回線3に送出する。圧縮された映像信号（以下では、映像データという）は通信回線3に接続されている映像サーバ4に一時的に蓄積される。ここで、映像サーバ4は、映像データのサイズや色調を携帯端末2の画面に表示可能なフォーマットに変換する変換手段としての機能を有する。したがって、携帯端末2から専用のインターネット接続サービス（携帯電話において提供されているインターネット接続サービス等）を通して映像サーバ4にアクセスすれば、映像サーバ4に蓄積された画像を携帯端末2の画面に表示させることができる。映像データは映像サーバ4によって携帯端末2の画面に表示可能なフォーマットに変換されているから、ロボット1で撮像された映像を携帯端末2によって確認することができる。なお、映像データは静止画でもよいが動画が望ましい。ただし、携帯端末2では一般に映像の記憶に用いるメモリ容量に制限があるから、映像データを携帯端末2に取り込んだ後に再生することは難しく、現状での通信回線3における通信速度および映像の圧縮技術を考慮するとテレビ電話程度の動画になる。

【0038】ところで、通信手段23には通知手段28が接続されており、当該ロボット1により撮像され映像サーバ4に蓄積された画像が携帯端末2によって閲覧されたときに、通知手段28は聴覚的手段（スピーカ）ないし視覚的手段（ランプ）を用いて対話者に画像が閲覧された旨を通知する。すなわち、対話者の家族などが携帯端末2を用いて映像サーバ4に蓄積されている画像を閲覧すれば、映像サーバ4の画像の閲覧がロボット1に通知され通知手段28により対話者に通知されることになる。このことによって、対話者は家族などが自身の状況を確認してくれたという安心感を持つことができ、不安感を解消することができる。

【0039】映像サーバ4に映像データを蓄積するタイミングは、携帯端末2からの指示によって決定する。つ

まり、携帯端末2からの指示により映像サーバ4を通してロボット1に映像データの蓄積を要求し、映像サーバ4に蓄積された画像を携帯端末2に転送させる。ただし、携帯端末2から映像データの転送を要求したときにロボット1と対話する対話者がロボット1の近くにいない場合があり、また映像データの転送を要求した時点よりも過去の映像データを閲覧したい場合もあるから、TVカメラ13による撮像をあらかじめ指定した間に定期的に行って映像データを映像サーバ4に蓄積しておいてもよい。このように、TVカメラ13で撮像した映像データを映像サーバ4に蓄積するタイミングと、携帯端末2によって映像サーバ4から映像データを読み出すタイミングとをずらせば、携帯端末2から映像サーバ4に指示を与えた後に、映像サーバ4から携帯端末2に映像データを転送する時間だけ待てば映像データを見ることができるから、映像を見るまでの応答時間を短縮することができる。なお、この場合には各映像データに撮像時刻を添付しておくのが望ましい。

【0040】また、映像サーバ4に蓄積された画像が他人によって閲覧されることがないように、映像サーバ4に蓄積された各映像データへのアクセスは、各携帯端末2からの識別情報に応じて制限される。たとえば、携帯端末2として携帯電話を用いる場合には、携帯電話の電話番号を識別情報に用い、映像サーバ4では当該電話番号に対応付けられている映像データのみを携帯電話に転送する。識別情報は電話番号に限定されるものではなく、映像サーバ4にアクセスした後に暗証番号を識別情報に用いたり、複数の情報の組み合わせを識別情報に用いたりすることもできる。

【0041】上述のように、映像サーバ4に蓄積された映像データにアクセスする際には、携帯端末2からの識別情報を確認するために、映像サーバ4には識別情報と名前のような個人データとを対応付けたデータベースを設けてある。データベースは携帯端末2からの識別情報が入力されたときに、個人データをロボット1に転送し、対話者に対して通知手段28を通して個人データを通知するようにしてある。この目的のために通知手段28は音声によって個人データを通知するか、あるいは別途に小型のディスプレイ装置を設けて個人データを表示する。この構成により、対話者は映像サーバ4にアクセスした人の確認ができ、対話者に対して安心感を与えることができる。

【0042】ところで、本実施形態では、携帯端末2を通して得られた映像のうちから必要なものを選択し電子メールに添付して適宜の送付先に転送可能としてある。つまり、映像サーバ4では携帯端末2によって指定された映像データにインデックスを付与し、この映像データを携帯端末2で入力された電子メールの添付書類として扱うことができるようにしてある。したがって、携帯端末2を通して得た映像により撮像された対話者に何らかの



異常が認められるような場合には、かかりつけの医者に映像データを添付した電子メールを送信することが可能になり、医師のアドバイスを受けるなどの適切な処置が可能になる。

【0043】上述のように、映像サーバ4を通して携帯端末2で再生することができる映像をTVカメラ13で撮像するから、TVカメラ13によって撮像するタイミングがロボット1と対話する対話者にとっては不都合なときもある。たとえば、身内といえども着替え中に撮像されるのはプライバシー上で好ましくない。そこで、ロボット1と対話する対話者にTVカメラ13での撮像中であることを報知するために、ロボット1には表示手段27を設けてある。表示手段27は発光ダイオードでよく、TVカメラ13での撮像中に点灯させることによって対話者に報知することができる。

【0044】ところで、対話処理手段21は、図2に示す構成を有し、音声による簡単な対話が可能になっている。マイクロホン11を通して入力された音声は音声認識部31に入力される。ロボット1の使用環境に鑑みれば周囲雑音もマイクロホン11を通して入力されていることが多いから、音声認識部31では、マイクロホン11から入力された音声の特徴パラメータを抽出し、音声の継続性や周波数分布に基づいて音声と雑音とを分離するとともに話者認識を行う。話者認識を行うのは、テレビジョンやラジオからの音声と対話者の音声との誤認を避けるためである。話者認識や雑音の分離に要する情報をここでは音響モデルM1と呼ぶことにする。また、音声認識部31においては、標準パターンとの照合によって単語認識を行い、音声をテキストデータTXT1に変換して出力する。ここに、単語認識を行う標準パターンの情報をここでは言語モデルM2と呼ぶことにする。

【0045】テキストデータTXT1は対話管理部32に入力される。対話管理部32は、一種の推論エンジンであって、テキストデータTXT1を知識ベースIBに格納された知識と照合する自然言語処理技術を用いて、テキストデータTXT1の構文解析やキーワードの抽出を行い、その結果に基づいて意味解析や文脈解析を行う。すなわち、対話者による発話の意味や対話者の意図を分析して所定の形式で表し、発話の意味や対話者の意図の分析結果に応じて応答内容を生成して所定の形式で発話内容生成部33に引き渡す。ここで、所定の形式とは意味表現の形式のことである。発話内容生成部33では、対話管理部32から渡された応答内容に応じてテキストデータTXT2を生成する。このテキストデータTXT2は、音声発生部34に引き渡され、テキストからの音声合成を行ったり、予め登録されている音声をテキストに従って編集することによって、スピーカ12を通して音声として出力させる。ただし、スピーカ12から出力する音声をできるだけ自然な音声に近付けることが望ましい。

【0046】対話管理部32においては、以下に説明するように、「しり取り」、「スケジュール管理」に対応した応答内容を生成することができる。どの応答内容を生成するかは、対話者がマイクロホン11を通して音声によってロボット1に伝達する。

【0047】「しり取り」を行うには、対話者が「しり取りがしたい」という命令を音声によってロボット1に与える。ここで、先攻と後攻とを取り決めた後、しり取りが開始される。しり取りでは、対話者の発話は原則として単語のみであるから、構文解析や文脈解析を行うことなく単語認識のみを行えばよいことになる。しり取りの際の処理手順は、図3のようになる。いま、対話者が先攻であるものとする、対話管理部32においては、音声認識部31からのテキストデータTXT1（単語）が入力されると（S1）、入力された単語がしり取りの開始後にすでに入力された単語か否かを判断する（S2）。しり取りの開始から未入力の単語であれば、適宜の記憶装置に記憶し、語尾の音節を抽出する（S3）。次に、抽出した語尾の音節と同じ音節の語頭を持つ単語を知識ベースIBの中の辞書から検索して抽出する（S4）。ここで、抽出した単語を上記した記憶装置に照合して、しり取りの開始後にすでに用いられた単語か否かを判定し（S5）、すでに用いられた単語であれば、ステップS4、S5を繰り返すことにより別の単語を検索して抽出する。ここに、辞書内での単語の並びに変化がなければ、毎回同じ順序で単語が抽出されることになって興味が失われるから、検索時には単語をランダムに検索するような検索方法が望ましい。こうして、適宜の単語が抽出されると、抽出した単語を上記した記憶装置に記憶するとともに、発話内容生成部33に対して抽出した単語の発話を指示し、次に音声認識部31から単語が入力されるのを待機する（S6）。ところで、ステップS2において音声認識部31から入力されたテキストデータTXT1を上記した記憶装置に照合し、しり取りの開始後にすでに用いられた単語と判断したときには、対話管理部32から発話内容生成部33に対して「前に言いましたよ」というテキストデータTXT2を生成するように指示を与える（S7）。この処理によって、対話者に別の単語の発話を促すことになる。しり取りの終了は、対話者が「もう止める」という命令を音声によってロボット1に与えればよい。また、しり取りを止めた時点で、上述した記憶装置の内容は消去される。

【0048】なお、上述の例は簡単な手順を示すものであり、語尾が「ん」の単語を言うとき負けというルールや、応答が遅いと負けというルールを含んだ処理になっていないが、これらのルールを実現する処理を含めておくのが望ましい。たとえば、対話者とロボット1との用いた単語の語尾を確認する処理を含めるのが望ましい。この場合、ロボット1側も辞書から抽出した単語の採用後に語尾の確認を行うようにする。また、上述の処

理ではロボット1は辞書に登録されている単語をすべて用いることができ、ロボット1が対話者に負ける可能性が少ないから、ロボット1の応答時間をランダムに制御して適宜に負けさせるなどの処理を行うのが望ましい。

【0049】以下に、「しり取り」の進行例を説明する。ここでは、対話者が「リング」と言い、ロボット1が「ゴリラ」と応答する場合を例にして説明する。対話者の音声は音声認識部31において認識され「リング」というテキストデータTXT1が生成されて対話管理部32に渡される。ここで、対話管理部32では、テキストデータTXT1の内容確認を行うように発話内容生成部33に指示する。つまり、発話内容生成部33には、「(テキストデータTXT1)ですね」という確認用の文言が定型文として登録しており、対話管理部32ではこの定型文を用いるように発話内容生成部33に指示を与える。この例では(テキストデータTXT1)=(リング)であるから、発話内容生成部33では「リングですね」というテキストデータTXT2を生成し、音声発生部34およびスピーカ12を通して対話者に応答する。ロボット1がスピーカ12から「リングですね」と発話したことに対して対話者が肯定の応答をすると(たとえば、「はい」「イエス」「うん」などと答える)、対話者の発話内容が「リング」であったと確認することができる。つまり、対話者の発話に対してロボット1が復唱して単語を確認するのである。ここに、音声認識部31において学習機能を持たせておけば、単語毎に復唱を繰り返すことで音声認識部31に対話者の発音の特徴パラメータを学習させることが可能になり、話者認識および単語認識の認識率を高めることができる。

【0050】対話者から肯定の応答があったときには、図3に示した手順の処理を行うことにより、対話管理部32では対話者が発話した単語の語尾から「ゴ」を抽出し、「ゴ」を語頭に持つ単語を知識ベースIBの辞書から抽出する。本例では「ゴリラ」が抽出されるものとしており、これによって対話管理部32は発話内容生成部33に対して「ゴリラ」のテキストデータTXT2を生成するように指示する。ここにおいて、上述した記憶装置には「リング」と「ゴリラ」との単語が登録される。こうして、ロボット1はスピーカ12から「ゴリラ」という単語の音声を出力する。以下同様にして、対話者とロボット1との一方が負けるか、対話者が終了を指示するまで上述の処理が繰り返される。

【0051】次に、「スケジュール管理」について説明する。「スケジュール管理」においては、音声認識部31から対話管理部32に与えられたテキストデータTXT1に、スケジュールに関するキーワードが含まれているか否かを判定することによって起動する。たとえば、「起きる」、「薬を飲む」などの概念を含むテキストデータTXT1が対話管理部32に入力されると、対話管理部32ではスケジュールに関する情報と認識する。こ

こでは、目覚ましの機能について例示する。この機能を実現するために、図1に図示していないが、ロボット1には現在日時を計時する時計手段51(図6参照)が設けられている。いま、対話者がロボット1に向かって「5時に起こして」と言ったとすると、対話管理部32では知識ベースIBを用いて構文解析や文脈解析を行うことによって、「5時に」を時刻と理解し、「起こして」を目覚ましのセットの指示と理解する。つまり、知識ベースIBには、「(数字)時に」の「(数字)」は時刻を表すという知識が設定されており、このような知識を参照して目覚ましのセットを行うのである。

【0052】ところで、音声認識部31における単語認識は複数の単語の候補が抽出されたときには類似度の高いものを抽出しており、構文や文脈による解析を行っていないから単語を誤認識することもある。たとえば、「5時に」を「工事に」と誤認識する場合もある。この場合、対話管理部32に入力されるテキストデータTXT1は、「工事に起こして」になる。対話管理部32においては入力されたテキストデータTXT1に対して構文や文脈から意味を解析しているから、知識ベースIBに照合することによって、「工事に」の次に「起こして」は結合できず意味の解析ができないと判断する。このような場合には、音声認識部31での誤認識と判断し、「起こして」の前に結合可能で「工事に」と音韻の類似している候補と、「工事に」の後に結合可能で「起こして」と音韻の類似している候補とを抽出し、意味が成立する候補を仮のテキストデータを生成する。いま、「工事に」を「5時に」に置き換えて「5時に起こして」という仮のテキストデータを生成したとする。このようにして生成したテキストデータは仮のものであるから、対話者に対して確認することによって検証する。つまり、「しり取り」の場合と同様にして、仮のテキストデータを確認するための質問文のテキストデータTXT2を生成するように発話内容生成部33に指示する。たとえば、仮のテキストデータが「5時に起こして」であるから、「5時に目覚ましをセットしますよ」という質問文を発話内容生成部33で生成し、このテキストデータTXT2を用いて対話者に音声での確認を行う。この質問に対して対話者から肯定の応答があれば、仮のテキストデータが正しかったものとして目覚ましを5時にセットする。また、仮のテキストデータに対して否定の応答であれば、「もう一度ゆっくり言って下さい」などの音声を出力する。このような対話によって対話者が音声によって簡単なスケジュール管理の指示をロボット1に与えることができる。

【0053】ここにおいて、対話管理部32での構文解析や意味解析で結合できない単語が結合されているときに、単語認識の際に誤認識した可能性があるものとして、仮のテキストデータを生成するとともに質問文を生成し音声によって確認する処理は、スケジュール管理だ

けではなく、通常の対話の際にも行われる。つまり、一種の復唱の処理がなされることになる。また、復唱して確認することにより、対話者の応答を教師信号に用いてロボット1を学習させることができ、対話者の音声に対する認識率を高めることができる。

【0054】本実施形態において、映像データを携帯端末2に表示可能なフォーマットに変換する変換手段としての機能を映像サーバ4に設けた例を示したが、ロボット1の映像処理手段22と通信手段23とのいずれかに、映像データを携帯端末2に表示可能なフォーマットに変換する変換手段としての機能を設けるようにしてもよい。また、通信手段23は必ずしもロボット1に内蔵していなくてもよく、通信回線23に接続される通信手段23とロボット1との間は宅内に無線通信路を形成することによって接続してもよい。

【0055】(第2の実施の形態)第1の実施の形態においては、ロボット1が犬型ではあっても犬様の動きを行うものではなかったが、本実施形態では、ロボット1が適宜の関節を持ち、図4に示すように、複数のアクチュエータからなるロボット駆動部24を制御して関節を屈伸ないし回転させることによって犬様の動作が可能になっている例について説明する。ロボット駆動部24はロボット制御手段25を通して信号処理部20に接続され、信号処理部20からの指示に応じてロボット駆動部24が制御される。ロボット制御手段25は、信号処理部20からの指示をロボット駆動部24の動作に変換する。すなわち、対話処理手段21の音声認識部31ではロボットに対する命令の語彙(単語)を抽出可能としておき、対話管理部32においてロボット1に対する命令であることを認識すると、命令内容の分析結果をロボット制御手段25に引き渡す。ロボット制御手段25では命令内容の分析結果に応じてロボット駆動部24の各アクチュエータの操作量をそれぞれ決定する。たとえば、ロボット1の移動について、「回れ右」、「回れ左」、「少し右」、「少し左」などの命令を認識可能としてあり、「少し」というような曖昧語に対しても対応可能としてある。

【0056】さらに、本実施形態では、ロボット1を移動させることなくTVカメラ13の光軸の向きや焦点距離を変化させることが可能にしてある。つまり、TVカメラ13を左右に振る機能とTVカメラ13に設けたズームレンズを駆動する機能とを有したカメラ制御手段26が設けられている。カメラ制御手段26に対する指示は信号処理部20を通して行われる。信号処理部20では、対話者の音声による指示を、カメラ制御手段26に与えることによって、TVカメラ13の視野を制御する。たとえば、対話者の音声によってTVカメラ13の視野を制御する場合には、ロボット1の移動を指示する場合と同様に、対話処理手段21において、「ズームイン」、「ズームアウト」、「少し右」、「少し左」など

の命令を認識可能としておき、これらの命令に応じてTVカメラ13の焦点距離や向きを制御するのである。なお、本発明におけるTVカメラ13は主として対話者を撮像することを目的としているから、映像処理手段22においてはTVカメラ13からの映像信号に基づいて人の顔の位置を認識し、ズームインにおいては人の胸から上が視野内に入る程度を最大とし、ズームアウトにおいては人の全身が視野内に入る程度を最小とするのが望ましい。また、「少し右」、「少し左」という命令に対しては、人の左右が視野内に入る程度が望ましい。

【0057】ロボット1の移動に関する命令およびTVカメラ1の制御に関する命令は、対話者の音声によって指示する以外にも、携帯端末2から通信回線3を通して与えることもできる。ただし、通信回線3を通して外部からロボット1に命令を与えるから、識別情報による認証作業が必要であって、たとえば暗証番号を用いることによって特定の人以外がロボット1に命令を与えられないようにしてある。携帯端末2からロボット1に対しては、番号の組み合わせによって命令を与えるようにすることが可能であるが、対話処理手段21を用いて携帯端末2からの音声を受け付けるようにすれば、マイクロホン11を用いて命令を与える技術をそのまま利用することができる。

【0058】ところで、ロボット1に対してマイクロホン11から命令を与える構成と携帯端末2から命令を与える構成とを併用すると、ロボット1が互いに矛盾する命令をほぼ同時に受けることがある。この場合、ロボット1が命令を順に処理したとしても、後に処理された命令に対応した動作を行うことになるから、先に処理された命令を与えた側から言えば、ロボット1が意図に反した動作をしたことになる。そこで、ロボット1に命令を与えるに際して、一方の命令が実行されると当面(命令に対する一通りの動作が完了する程度の時間)は他方の命令を受け付けられないようにし、他方の命令を受け付けられない期間においては当該他方に対して上記一方の命令の実行中であることを報知するようにしてある。たとえば、ロボット1と対話する対話者が音声による命令を与えようとするときに携帯端末2からの命令を実行中であれば、「携帯端末からの操作中」である旨の音声メッセージをスピーカ12を通して出力する。逆に、携帯端末2を操作する対話者が命令を与えようとするときにロボット1と対話する対話者が音声によって与えた命令を実行中であれば、「対話による操作中」である旨のメッセージを携帯端末2の画面に表示する。携帯端末2からロボット1に対して音声による命令を与える場合であれば、「対話による操作中」である旨の音声メッセージを携帯端末2を通して出力してもよい。このように、対話者と携帯端末2とのうちの一方からの指示によってロボット1が動作しているときに、命令の被指示側である他方に対して命令が受け付けられない旨の報知を行うように、

信号処理部20には図示しない報知手段が設けられる。

【0059】ところで、本実施形態では、対話者からの情報を得るインタフェースとしてマイクロホン11とTVカメラ13とのほかに接触センサ14を備えている。接触センサ14はロボット1の表面付近の適宜箇所に配置され、接触センサ14の出力はマイクロホン11やTVカメラ13の出力と同様に信号処理部20に入力される。しかして、接触センサ14により対話者がロボット1に触れたことが検出されると、対話処理手段21を起動するようにしてある。つまり、対話処理手段21を常時動作させていると、周囲の雑音に反応してロボット1が動作することがあるから、対話処理手段21は対話者が発話している期間にのみ起動するのが望ましいのであるが、対話者がいつ発話するかは予測できない。周囲の雑音よりも大きい音声はマイクロホン11に入力されると対話処理手段21を起動する技術も考えられるが、このような技術を採用すると語頭が切断され、発話内容を誤認識する可能性が高くなる。また、対話処理手段21において話者認識の技術を用いても、音声のみで話者を完全に特定することは現状の技術では難しいものであるから、周囲の雑音を話者と誤認する可能性もある。

【0060】そこで、本実施形態においては、対話者がロボット1に触れたことが接触センサ14に検出されると対話処理手段21を起動するのであって、対話者以外の音声にロボット1が反応する可能性をほぼ回避できることになる。

【0061】本実施形態では、第1の実施の形態と同様に、携帯端末2から映像サーバ4に蓄積された映像が閲覧されたときに個人データが転送される。ただし、通話手段28には個人データにロボット1の動きを対応付けるテーブルが設けてあり、ロボット1が図5のような犬型であるとすれば、頭部1aや尾部1bに各個人データを対応付ける。つまり、個人データに応じて、頭部1aを振る動作や尾部1bを振る動作に対応付けてある。このことによって、対話者は映像サーバ4にアクセスしたのが誰かを直感的に知ることが可能になる。なお、個人データを識別するための報知部として、上述のようにロボット1の各部の動作を用いるほか、個人データに色を対応付けた表示灯を報知部として用いることも可能である。

【0062】ところで、本実施形態におけるロボット1は、インターネットへの接続が前提とされており、通信手段23が、電子メール処理部41とインターネット検索処理部42とを備えている。インターネットを通して電子メールを交換したり情報を検索したりする処理は、「しり取り」や「スケジュール管理」と同様に、対話者がマイクロホン11を通して音声によってロボット1に指示することができる。

【0063】いま、携帯端末2との間で電子メールを交換する場合について例示する。ここに、送信先のメール

アドレスは電子メール処理部41に設定されているものとする。ロボット1と対話する対話者から携帯端末2に電子メールを送信する場合は、対話者がロボット1に対して「電子メールを××に送信したい」と音声によって伝えることで、電子メール処理部41が起動し、電子メールの本文を入力できるようになる。ここで、対話者が音声によって電子メールの本文を作成し、「電子メールを送信して」というように音声によってロボット1に命令を与えると、ロボット1が本文を読み上げて内容の確認を行い、対話者が承諾すると、作成された電子メールがインターネットプロバイダのメールサーバに蓄積される。メールサーバに蓄積された電子メールは、通常の電子メールと同様であるから、携帯端末2によって読み出すことができる。

【0064】また、ロボット1と対話する対話者に宛てた電子メールをメールサーバから読み出すためにロボット1はメールサーバに定期的にアクセスしており、また対話者が音声によって電子メールの読み出しをロボット1に指示した場合にも電子メールの読み出しが可能になる。電子メール処理部41においてメールサーバから電子メールが読み出されると、電子メールの本文はテキストデータTXT2として音声発生部34に入力される。したがって、電子メールの本文が読み上げられ、音声によって対話者に伝えられることになる。

【0065】インターネット検索処理部42はキーワード検索に用いられる。インターネット検索処理部42は、あらかじめ設定された検索用のホームページに接続する機能を有し、そのホームページは検索エンジンを持ち、検索結果をテキストデータとして転送できるようにしてある。つまり、ロボット1に専用の検索用ホームページが設けてある。この検索用ホームページは、基本的には電話帳の機能を有するものであり、対話者が希望する地区で希望する業種の業者を検索する機能を有している。たとえば、対話者が「宅配業者を探して」と音声によってロボット1に伝え、対話管理部32においてインターネットでの検索が要求されていると認識して、インターネット検索処理部42を通して宅配業者を検索する。検索結果はテキストデータTXT2として音声発生部34に与えられ、検索された宅配業者名が音声によって順に読み上げられる。ここで、検索結果の数が多いときにはインターネット検索処理部42が絞り込みを促すように発話内容生成部33に指示し、検索結果が適当数になれば、検索結果が音声によって対話者に伝達される。さらに、対話者が希望する宅配業者を選んで音声によってロボット1に伝え、その宅配業者の電話番号がホームページから読み出されて音声により対話者に伝達される。

【0066】上述した各実施形態においては、対話者がロボット1に指示を与えることによってロボット1が動作するようにしているが、従来構成として説明したペットロボットでは、自律エージェントによってロボット1

が自動的に起動して対話者に話しかける機能を持ち、また対話者の表情や身振りなどの情報も用いて音声認識を行うマルチモーダル対話の機能も備えているものであって、本発明においてもこれらの機能を付加することによって、対話者とロボット1との間でよりスムーズな対話が可能になる。他の構成および動作は第1の実施の形態と同様である。

【0067】(第3の実施の形態) 本実施形態は、第1の実施の形態において、より利便性を高めるように機能を付加したものである。本実施形態において付加する機能は第2の実施の形態においても適用可能である。

【0068】本実施形態では、図6に示すように、TVカメラ13により撮像された映像に撮像された日時を付加するために、現在の日時を計時する時計手段51と、時計手段51により計時された日時をTVカメラ13で撮像された映像に文字としてスーパーインポーズ表示するためのスーパーインポーズ処理手段52とを備える。すなわち、TVカメラ13により図7(a)のような映像が撮像され、時計手段51により計時されている撮像日時が2000年11月15日の15時23分であったとすると、図7(b)のように、TVカメラ13で撮像された映像の上部あるいは下部に、「2000/11/15 15:23」などの形式で映像の撮像日時を示す文字が合成される。このような日時付きの映像が映像サーバ4に転送されることによって、携帯端末2の画面にもTVカメラ13で映像が撮像された日時が表示されることになる。つまり、対話者がTVカメラ13で撮像された時点から携帯端末2の画面に映像が表示されて対話者の様子が確認されるまでには時間差があるから、携帯端末2の画面に映像の撮像された日時を表示することによって、少なくともその日時での対話者の様子を保証することが可能になる。なお、スーパーインポーズ処理手段52は、マイコンよりなる映像処理手段22に適宜プログラムを与えて実現してもよいが、専用の集積回路によって構成すれば映像処理手段22における負荷が低減される。

【0069】ところで、映像サーバ4における記憶容量は非常に大きいものではあるが、ロボット1の台数が多くなれば映像データのデータ量が膨大になるから、映像サーバ4ではTVカメラ13で撮像されたすべての映像データを蓄積するのではなく、携帯端末2で閲覧された映像データは原則として消去することになる。一方、対話者に何らかの異常が生じた場合であって異常の発生時点から携帯端末2で映像が閲覧されるまでの間に大きな時間差があったような場合には、異常の発生前から発生後に跨る期間の映像データの履歴が要求されることがある。上述のように映像サーバ4に蓄積された映像データは逐次消去されるから、このような要求に十分に対応することができない場合がある。また、通信手段23を通してロボット1から映像サーバ4に映像を転送するもの

であるから、通信回線3の状態や映像サーバ4の状態によっては、映像サーバ4に映像データが蓄積されない場合も考えられる。

【0070】そこで、本実施形態では通信手段23を通して映像サーバ4に転送する映像データを蓄積するための記憶手段53をロボット1に付加してある。記憶手段53は電源が遮断された状態でも記憶内容を保持することができるものを用いている。この種の記憶手段53としては、フレキシブルディスク、ハードディスク、CD-R、不揮発性メモリなどの各種のものを用いることができる。記憶手段53に格納される情報は、各通信手段23から送出される映像データと、各映像データの送信の日時および送信の成功・失敗のデータとが含まれる。たとえば、送信する映像データ毎に個別の名称を付し、「送信日時、映像データの名称、送信の成功・失敗」というような形式のデータを映像データとともに蓄積すればよい。また、映像データに付ける個別の名称は送信日時を利用して設定すればよい。たとえば、送信日時が2000年11月13日8時13分であるとすれば、映像データの名称を「200011130811a」などとすればよい。このように、通信手段23から送出する映像データと送信の履歴とをロボット1に蓄積することによって、必要に応じて映像データを過去に遡って閲覧することが可能になる。ここに、記憶手段53に蓄積された映像データおよび送信履歴については、記憶手段53から記録媒体を外すことによって別のパーソナルコンピュータなどによって読み出すことが可能である。

【0071】上述の説明は対話者がTVカメラ13で確実に撮像されていることを前提にしているが、実際にはロボット1が傾いている場合や、対話者がTVカメラ13の視野の中心付近に存在しない場合もある。そこで、以下では対話者をTVカメラ13で適正に撮像する技術について説明する。

【0072】ロボット1が傾いて設置されている場合や対話者がロボット1を抱いているとすると、TVカメラ13により撮像された映像が傾くことがある。そこで、本実施形態においては、ロボット1の傾きを検出するための姿勢センサ54を付加してある。姿勢センサ54としてはジャイロセンサを用いることができる。姿勢センサ54によって検出されたロボット1の傾きは映像処理手段22に与えられ、映像処理手段22において映像が正立するように映像の傾きが補正される。ここに、映像が正立するとは重力方向に直交する面が映像の下縁に平行になる状態を意味する。また、映像の左右の傾きを補正すればよく、上下の傾きについては必ずしも補正しなくてもよいが、上下左右の傾きをともに補正すればTVカメラ13を水平面上に設置している状態と同様の映像を得ることが可能になる。たとえば、姿勢センサ54により検出されたロボット1の傾きによって映像が左回りに30度回転していると判断されると、映像処理手段2

2においては映像を画面の中心を中心として右回りに30度回転させる補正を行い、補正後の映像を用いて以後の処理を行うのである。

【0073】また、TVカメラ13の視野内に対話者を捕捉するために、TVカメラ13の向きを可変とする首振り駆動手段55と、TVカメラ13により撮像すべき所要の映像をあらかじめテンプレートとして登録しておくTVカメラ13により撮像している映像をテンプレートと照合したときの類似度が大きくなるように首振り駆動手段55を制御するパターン処理手段56とを設けている。首振り駆動手段55はTVカメラ13の光軸の向きを変化させるように設けられ、たとえばTVカメラ13をロボット1の頭部に配置しているとすれば首振り駆動手段55によりロボット1の頭部の位置ないし向きを移動可能とすればよい。また、ロボット1はとくに移動させずにロボット1の本体に対してTVカメラ13のみが移動可能となるように首振り駆動手段56を構成してもよい。

【0074】一方、パターン処理手段56はTVカメラ13により撮像された映像をテンプレートとして登録されている映像と照合し、得られている画像の範囲で類似度が最大になる方向に首振り駆動手段55を駆動してTVカメラ13の光軸の向きを変更する。たとえば、テンプレートとして図8(a)のような映像を登録しているときには、図8(b)や図8(d)のような映像よりも図8(c)の映像のほうが類似度が大きいから、首振り駆動手段55では図8(c)のような映像が得られるようにTVカメラ13の向きを変更する。このようにパターン処理手段56に、対話者の適切なテンプレートを登録しておけば、首振り駆動手段55を適正に制御して対話者をTVカメラ13でつねに捕捉することが可能になる。なお、対話者は同じ姿勢を保つわけではないから、形状や移動速度などに関する特徴抽出を行い、人か否かを判断する知識をテンプレートと併せて用いることによって首振り駆動手段55を制御するようにしてもよい。

【0075】上述したパターン処理手段56ではTVカメラ13により撮像した映像内に対話者が存在しなければ、首振り駆動手段55をどのように駆動すべきか判断することができない。そこで、マイクロホン11として指向性の強いものを用いて音レベルによっても対話者の存在する方向を知ることが可能としてある。つまり、本実施形態ではマイクロホン11が第2の音声入力装置として機能する。また、マイクロホン11における指向性が最大になる向きはTVカメラ13の光軸の向きと平行に設定され、首振り駆動手段55はマイクロホン11とTVカメラ13との位置関係を固定したままでマイクロホン11およびTVカメラ13の向きを変更するように構成してある。マイクロホン11から入力された音声の音レベルは音レベル処理手段57により検出され、パターン処理手段56において類似度が所定値以下であると

ときには、音レベル処理手段57で検出した音レベルが大きくなる向きに首振り駆動手段55を駆動するように構成されている。要するに、首振り駆動手段55によってマイクロホン11の向きを変更している間に、図9

(a)(c)に示すような音レベルの比較的小さい状態から図9(b)のように音レベルが比較的大きい状態に変化したとすれば、図9(b)の状態ではマイクロホン11の指向性が最大になる向きに対話者が存在しているとみなすことができる。

【0076】たとえば、対話者の音声はマイクロホン11を通して入力されると、対話者の音声が入力されている間に首振り駆動手段55によってマイクロホン11の向きが変更され、マイクロホン11に入力される音声の音レベルがほぼ最大になる向きにマイクロホン11が向けられる。このとき、TVカメラ13の光軸も同じ向きを向いており、マイクロホン11に入力される音声の音レベルが大きくなる向きは対話者が存在する向きである可能性が高いから、結果的にTVカメラ13の視野内に対話者を捕捉できる可能性が高くなるのである。なお、マイクロホン11の指向性を利用して対話者を捕捉する技術を用いる場合に、パターン処理手段56を省略することも可能である。また、対話者を捕捉するために用いる指向性の強いマイクロホンは、対話者との対話に用いるマイクロホン11とは別に設けることも可能である。

【0077】さらに、図10に示すように、TVカメラ13の前方においてTVカメラ13の光軸Axに直交させた形でハーフミラー58を配置してもよい。このようにTVカメラ13の前方にハーフミラー58を配置しておけば、ハーフミラー58に対話者自身の姿を映した状態とすることによって、TVカメラ13の視野内に対話者が確実に入ることになり、TVカメラ13の視野内に対話者が捕捉されていることを保証することができる。要するに、対話者の姿はハーフミラー58に反射されてハーフミラー58に映り、ハーフミラー58を透過した光によってTVカメラ13で対話者を撮像することができるのである。

【0078】対話者がTVカメラ13の視野内に存在することを対話者自身で確認するために、映像処理手段22で扱う映像信号を通常のテレビ受像機に出力可能とする映像出力手段59を付加してもよい。この場合には、映像出力手段59にテレビ受像機を接続しておけば、対話者がTVカメラ13により撮像されていることを対話者自身がテレビ受像機の画面によって確認することができる。また、対話者は自身の姿のみではなく、室内の様子など周辺の様子もテレビ受像機の画面によってその場で認識することが可能になる。つまり、テレビ受像機を映像確認用のディスプレイ装置として用いることが可能になる。また、映像出力手段59は必ずしもテレビ受像機の画面に映像の表示を可能とするものでなくてもよく、TVカメラ13で撮像している映像が表示可能なデ



ディスプレイ装置を備えるものであれば、パーソナルコンピュータのディスプレイ装置などに映像を出力するようにすることも可能である。ディスプレイ装置に出力された映像が対話者にとって不満である場合に備えて対話者によって当該映像を消去可能とする機能をロボット1に設けておくことが望ましい。この機能については音声でロボット1に命令できるようにしておくのが望ましい。

【0079】上述したように、ロボット1は犬型の外観を呈するものであり、いわゆるペットロボットとして用いることができる。そこで、対話者がロボット1を抱けるように、ぬいぐるみのような外殻としかつ軽量に形成しておくことによって、必要に応じて対話者にロボット1を抱かせることが可能になる。対話者にロボット1を抱かせれば、ロボット1が対話者に接触することにより、音声や映像だけではなく接触感覚によっても対話者に関する情報を得ることが可能になる。そこで、本実施形態では心拍音を検出可能とする心拍センサとしての心拍用マイクロホン60をロボット1の胸部ないし腹部に配置してあり、対話者がロボット1を抱いたときに心拍用マイクロホン60により心拍音を集音するようにしてある。

【0080】心拍用マイクロホン60の出力は、たとえば図11のような波形になるから、心拍用マイクロホン60の出力についてサンプリングを行い、波形に関する特徴抽出を行うことによって心拍数や心拍波形の異常の有無などを検出することが可能になる。ただし、心拍用マイクロホン60は、必ずしも心拍音のみを検出するわけではなく、心拍用マイクロホン60の出力には雑音成分が多く含まれるから、適宜のフィルタリングが必要になる。

【0081】心拍用マイクロホン60の出力に基づいて得られた情報は、たとえば心拍数を表す文字を映像内にスーパーインポーズ表示して映像サーバ4に転送される。あるいはまた、心拍用マイクロホン60で集音した心拍音を圧縮し、映像データに添付して映像サーバ4に転送したり記憶手段53に記憶させるようにしてもよい。これによって、対話者の健康状態の目安を得ることが可能になる。記憶手段53に心拍音を記録させる場合には記憶手段53が心拍記録装置として兼用されることになる。また、記憶手段53とは別に心拍音を記録する心拍記録装置を設けてもよい。なお、対話者がロボット1を抱かなければ心拍音を集音することができないから、ロボット1からの音声メッセージによってロボット1を抱くように指示させるのが望ましい。このような音声メッセージは所定時刻毎に発生させるのが望ましく、たとえば起床時刻の直後において日々の心拍音を検出するようにすれば健康状態の管理に役立つことになる。

【0082】(第4の実施の形態)第3の実施の形態においては、対話者をTVカメラ13の視野内に捕捉するためにマイクロホン11として指向性を有するものを用

い、1個のマイクロホン11で検出される音声の音レベルが大きくなる向きに首振り駆動手段55を制御していたが、本実施形態では図12に示すように、2個のマイクロホン11a, 11bを用い両マイクロホン11a, 11bで検出される位相差に基づいて首振り駆動手段55を制御するように構成してある。つまり、マイクロホン11a, 11bは第3の音声入力装置として機能する。2個のマイクロホン11a, 11bはロボット1の左右の耳に対応させて配置してあり、音源である対話者と各マイクロホン11a, 11bとの距離が等しいときに、図13(b)のように両マイクロホン11a, 11bの出力の位相差がゼロになるようにしてある。また、両マイクロホン11a, 11bの出力の位相差がゼロになる位置に対話者が存在するときにTVカメラ13の視野の中央に対話者が位置するように、両マイクロホン11a, 11bとTVカメラ13との位置関係を設定してある。

【0083】したがって、対話者がTVカメラ13の視野に対して左側に位置するときには、図13(a)のように左側のマイクロホン11aの出力の位相が右側のマイクロホン11bの出力の位相よりも進むことになり、逆に対話者がTVカメラ13の視野に対して右側に存在するときには、図13(c)のように右側のマイクロホン11aの出力の位相が左側のマイクロホン11bの出力の位相よりも進むことになる。そこで、両マイクロホン11a, 11bの出力の位相差を検出する位相差検出手段61を設け、位相差が小さくなる向きに首振り駆動手段55を制御することによって、対話者がTVカメラ13の視野の中央付近に位置するようにTVカメラ13の向きを調節することが可能になるのである。

【0084】さらに、本実施形態では2個のマイクロホン11a, 11bの出力の位相差から音源が存在する向きを求めているが、3個以上のマイクロホンの位相差を用いると音源の存在位置をさらに正確に把握することが可能になる。本実施形態の他の構成および動作は上述した他の実施の形態と同様である。

【0085】

【発明の効果】請求項1の発明は、音声入力装置と音声出力装置とが接続され自然言語を用いた音声による対話を対話者との間で行う対話処理手段と、対話者を撮像する映像入力装置が接続され映像入力装置で撮像された映像をデジタル化した映像データを出力する映像処理手段と、映像データをネットワークに送出する通信手段と、ネットワークを通して対話者の映像データの閲覧が許可されている携帯端末と、映像データを携帯端末に表示可能なフォーマットに変換する変換手段とを備えるものであり、対話処理手段を用いて人との対話を模擬するから対話者は孤独感に対する精神的なケアがなされる。また、対話者の映像データの閲覧は許可されている携帯端末のみで可能であるから、携帯端末を対話者の家族のみ

が携帯できるようにするように使用することによって、他人のプライバシーを侵害しない範囲で対話者の画像を携帯端末に転送して対話者の異常の有無を監視可能とすることができる。その結果、高齢者や病人を抱える家庭においても高齢者や病人の様子を携帯端末により遠方から監視することができ、安心して外出することができるようになる。

【0086】請求項2の発明は、請求項1の発明において、前記変換手段がネットワークを介して転送された映像データを蓄積する映像サーバに設けられ、映像サーバでは携帯端末の所持者に許容される映像データのみを選択して携帯端末に転送するものであり、対話者の映像データを蓄積する映像サーバを設けたことによって、過去に遡って映像データを閲覧することが可能になる。

【0087】請求項3の発明は、請求項1の発明において、前記変換手段が映像処理手段と通信手段との一方に設けられているものであり、映像サーバを設けることなく対話者の映像データを携帯端末から閲覧することができ、システム構成が簡単である。

【0088】請求項4の発明は、請求項1ないし請求項3の発明において、前記映像処理手段が前記携帯端末からの指示により対話者の映像データを通信手段を通してネットワークに送出させるものであり、携帯端末から映像データの送出を指示するから、携帯端末の所持者が要求したときの映像を見ることができる。

【0089】請求項5の発明は、請求項1ないし請求項4の発明において、前記映像処理手段が前記対話処理手段を通して対話者が指示した命令に回答して動作するものであり、対話者が自身の意思で映像を撮影させるから、たとえば対話者が異常を感じたときに映像データを送って携帯端末の所持者に知らせることが可能になる。

【0090】請求項6の発明は、請求項1ないし請求項5の発明において、前記映像処理手段が対話者からの指示により動作しているか前記携帯端末からの指示によって動作しているかを被指示側に報知する報知手段を備えるものであり、対話者からと携帯端末からとのどちらからも映像処理手段の制御を可能としながらも、両者の指示が混乱することがない。

【0091】請求項7の発明は、請求項1ないし請求項6の発明において、対話者の接触を検出する接触センサが付加され、接触センサにより対話者の接触が検出されると前記対話処理手段が起動するものであり、対話者が対話しようとするときに接触センサに接触することによって対話処理手段が起動されるから、周囲の雑音や対話者以外の他人の音声によって対話処理手段が誤動作するのを防止することができる。

【0092】請求項8の発明は、請求項1ないし請求項7の発明において、前記映像入力手段によって対話者の映像が撮影されていることを対話者に示す手段を備えるものであり、対話者がいつ撮影されているかを知ること

ができるから、プライバシーの都合で撮影されたくないときには撮影されない場所に移動するなどの対処が可能になる。

【0093】請求項9の発明は、請求項1ないし請求項8の発明において、前記対話処理手段が、対話者による音声入力を確認するように対話者に質問する機能を有し、対話者による肯定の応答を受けると対話者による音声入力を確定するものであり、復唱によって対話者の音声入力を確認するから、対話者の発生した音声を確実に理解することが可能になる。

【0094】請求項10の発明は、請求項9の発明において、前記対話処理手段が、対話者による肯定の応答を教師信号として対話者の発生する音声を学習するものであり、対話者との対話によって対話者の発生する音声を学習することで認識率が高くなる。

【0095】請求項11の発明は、請求項1ないし請求項10の発明において、前記携帯端末で閲覧した映像データを前記ネットワーク上の他の端末に転送する手段を設けたものであり、映像データを他の端末に転送することにより、たとえば携帯端末の所持者が対話者の異常に気づいたときに映像データを医者に送ることで対処のアドバイスを受けるなどの使用が可能になる。

【0096】請求項12の発明は、請求項1ないし請求項11の発明において、前記携帯端末から映像データの閲覧が要求されたことを前記対話者に通知する通知手段を備えるので、対話者である高齢者や病人を抱える家庭においても家族が携帯端末を用いて外出先から高齢者や病人の様子を監視したときに、高齢者や病人に対しては通知手段によって家族が確認したことが通知されるから、高齢者や病人にとっては家族が確認してくれたことを知ることができ安心感が得られる。

【0097】請求項13の発明は、請求項12の発明において、前記通知手段が前記携帯端末の識別情報を用いて映像データの閲覧者を通知する機能を備えるので、映像データを誰が確認したかを知ることができるから、より安心することができる。

【0098】請求項14の発明は、請求項13の発明において、前記通知手段が映像データの閲覧者ごとに割り当てた報知部を備えるので、文字データなどによらずに報知部を視認すれば閲覧者を識別することができるから、閲覧者が映像データを確認したことを知るだけでなく誰が閲覧したかも直感的に知ることができる。

【0099】請求項15の発明は、音声入力装置と音声出力装置とが接続され自然言語を用いた音声による対話を対話者との間で行う対話処理手段と、対話者を撮像する映像入力装置が接続され映像入力装置で撮像された映像をデジタル化した映像データを出力する映像処理手段と、映像データをネットワークに送出する通信手段と、ネットワークを通して対話者の映像データの閲覧が許可されている携帯端末に表示可能なフォーマットとなるよ

うに映像データを変換する変換手段とを備えるものであり、対話処理手段を用いて人との対話を模擬するから対話者は孤独感に対する精神的なケアがなされる。また、対話者の映像データの閲覧は許可されている携帯端末のみで可能であるから、携帯端末を対話者の家族のみが携帯できるようにするように使用することによって、他人のプライバシーを侵害しない範囲で対話者の画像を携帯端末に転送して対話者の異常の有無を監視可能とすることができる。その結果、高齢者や病人を抱える家庭においても高齢者や病人の様子を携帯端末により遠方から監視することができ、安心して外出することができるようになる。

【0100】請求項16の発明は、請求項15の発明において、現在日時を計時する時計手段と、対話者が撮像されたときに時計手段で計時されている日時を文字として映像データに合成して通信手段から送出させるスーパーインボーズ処理手段とが付加されたものであり、各映像データに映像の撮像された日時が添付されているから、映像が撮像された日時と映像データを携帯端末によって閲覧する日時とに大きな時間差があるようなときに、閲覧者は映像データに添付されている日時を確認することによって、最新の映像を求めるなどの対処が可能になる。

【0101】請求項17の発明は、請求項16の発明において、前記通信手段を通してネットワークに送出される映像データおよび通信の成功・失敗の結果を前記時計手段で計時されている日時に対応付けて順次記録して蓄積するとともに無給電で記録内容を保持する記憶手段が付加されているものであり、通信手段から映像データが送出されず携帯端末によって映像データを確認することができないような場合にも、記憶手段に蓄積された映像データによって対話者の健康状態などを後日確認することができ、また通信の成功・失敗の結果を送信履歴として記憶手段に格納しているから、メンテナンス時に通信手段の異常の有無などを容易に知ることができる。

【0102】請求項18の発明は、請求項15ないし請求項17の発明において、前記映像入力装置の傾きを検出する姿勢センサが付加され、前記映像処理手段では映像入力装置により撮像された映像を正立させるように姿勢センサにより検出された傾きに基づいて映像を補正するものであり、姿勢センサにより得られている情報に基づいて映像の傾きを除去するように補正するから、映像入力装置が傾いているような場合、たとえば水平ではない場所に撮影装置が配置されている場合や対話者が撮影装置を持ち上げているような場合においても、撮影装置が水平面上に配置されている状態と同様の映像を得ることができる。

【0103】請求項19の発明は、請求項15ないし請求項18の発明において、前記映像入力装置の視野の方向を変更可能とする首振り駆動手段と、前記映像入力装

置の中央付近に対話者を撮像するように規定したテンプレートが登録可能であって、映像入力装置で撮像された映像とテンプレートとを照合して類似度が大きくなるように首振り駆動手段を制御するパターン処理手段とが付加されているものであり、映像入力装置により撮像された映像とテンプレートとの類似度が大きくなるように映像入力装置の視野の向きが制御されるから、適切なテンプレートが設定されていれば対話者を映像入力装置の視野の中央付近に捕捉するように映像入力装置の向きを調節することが可能になる。

【0104】請求項20の発明は、請求項15ないし請求項18の発明において、前記映像入力装置の視野の方向を指向性を有する方向に一致させた第2の音声入力装置と、前記映像入力装置の視野の方向と第2の音声入力装置が指向性を有する方向とを同時に変更可能とする首振り駆動手段と、前記第2の音声入力装置の出力の音レベルが大きくなる向きに首振り駆動手段を制御する音レベル処理手段とが付加されているものであり、音レベルによって対話者の存在する方向を推定することができるから、映像入力装置の視野を音レベルの大きくなる向きに調節することによって、映像入力装置の視野内に対話者を捕捉できる可能性が高くなる。

【0105】請求項21の発明は、請求項15ないし請求項18の発明において、前記映像入力装置の視野の方向に存在する音源に対して出力の位相差がゼロになるように配置した複数の第3の音声入力装置と、第3の音声入力装置の位相差を小さくする向きに首振り駆動手段を制御する位相差検出手段とを付加したものであり、複数の音声入力装置の位相差を小さくするように映像入力装置の視野を調節することによって、映像入力装置の視野内に対話者を捕捉することが可能になる。

【0106】請求項22の発明は、請求項15ないし請求項21の発明において、前記映像入力装置の前方に前記映像入力装置の光軸方向に直交させた形で配置したハーフミラーが付加されているものであり、対話者がハーフミラーに自身の姿を映した状態では映像入力装置の正面に対話者が存在することになるから、対話者の意思によって対話者を映像入力装置で確実に撮像することが可能になる。

【0107】請求項23の発明は、請求項15ないし請求項21の発明において、外部のディスプレイ装置が接続可能であって前記映像入力装置により撮像された映像データをディスプレイ装置に表示可能とする映像出力手段が付加されているものであり、対話者は、映像入力装置により撮像されている映像を映像出力手段に接続したディスプレイ装置によってその場で確認できるから、対話者は自身の映り具合を確認することができるだけでなく周囲の様子も確認して適切な映像が撮像されているか否かを確認することができる。

【0108】請求項24の発明は、請求項15ないし請

求項23の発明において、前記対話者が抱えることを可能とする外殻を有し、前記対話者が抱えたときに対話者の心拍音を検出可能な心拍センサと、心拍センサにより検出された心拍音を記録する心拍記録装置とが付加されているものであり、映像によって対話者の健康状態を知るだけでなく、心拍音によっても対話者の健康状態を確認することが可能になり、対話者に関して得られる情報を複合することによって対話者の健康状態をより正確に知ることが可能になる。しかも、心拍センサとして対話者が抱えたときに心拍音を集音する構成を採用しているから、手間をかけることなく心拍音の集音が可能になる。

【0109】請求項25の発明は、請求項15ないし請求項24の発明において、前記携帯端末から映像データの閲覧が要求されたことを前記対話者に通知する通知手段を備えるので、話者である高齢者や病人を抱える家庭においても家族が携帯端末を用いて外出先から高齢者や病人の様子を看視したときに、高齢者や病人に対しては通知手段によって家族が確認したことが通知されるから、高齢者や病人にとっては家族が確認してくれたことを知ることができ安心感が得られる。

【0110】請求項26の発明は、請求項25の発明において、前記通知手段が前記携帯端末の識別情報を用いて映像データの閲覧者を通知する機能を備えるので、映像データを誰が確認したかを知ることができるから、より安心することができる。

【0111】請求項27の発明は、請求項26の発明において、前記通知手段が映像データの閲覧者ごとに割り当てた報知部を備えるので、文字データなどによらずに報知部を視認すれば閲覧者を識別することができるから、閲覧者が映像データを確認したことを知るだけでなく誰が閲覧したかも直感的に知ることができる。

#### 【図面の簡単な説明】

【図1】本発明の第1の実施の形態を示すブロック図である。

【図2】同上に用いる対話処理手段を示すブロック図である。

【図3】同上の対話処理手段の動作例を示す動作説明図である。

【図4】本発明の第2の実施の形態を示すブロック図で

ある。

【図5】同上に用いるロボットの外観を示す斜視図である。

【図6】本発明の第3の実施の形態を示すブロック図である。

【図7】同上の動作説明図である。

【図8】同上の動作説明図である。

【図9】同上の動作説明図である。

【図10】同上の動作説明図である。

【図11】同上の要部構成図である。

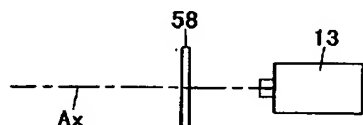
【図12】本発明の第4の実施の形態を示すブロック図である。

【図13】同上の動作説明図である。

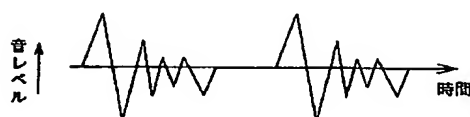
#### 【符号の説明】

- 1 ロボット
- 2 携帯端末
- 3 映像サーバ
- 4 通信回線
- 11 マイクロホン
- 11a, 11b マイクロホン
- 12 スピーカ
- 13 TVカメラ
- 14 接触センサ
- 20 信号処理部
- 21 対話処理手段
- 22 映像処理手段
- 23 通信手段
- 27 表示手段
- 28 通知手段
- 51 時計手段
- 52 スーパーバイポーズ処理手段
- 53 記憶手段
- 54 姿勢センサ
- 55 首振り駆動手段
- 56 パターン処理手段
- 57 音レベル処理手段
- 58 ハーフミラー
- 59 映像出力手段
- 60 心拍用マイクロホン
- 61 位相差検出手段

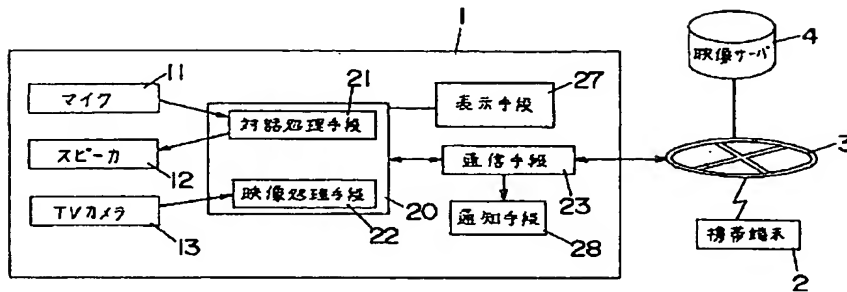
【図10】



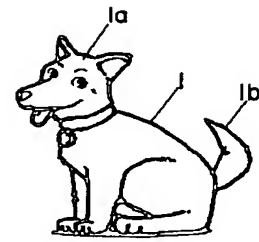
【図11】



【図1】

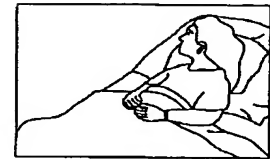


【図5】

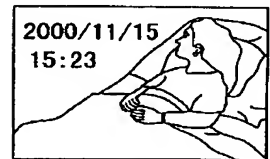


【図7】

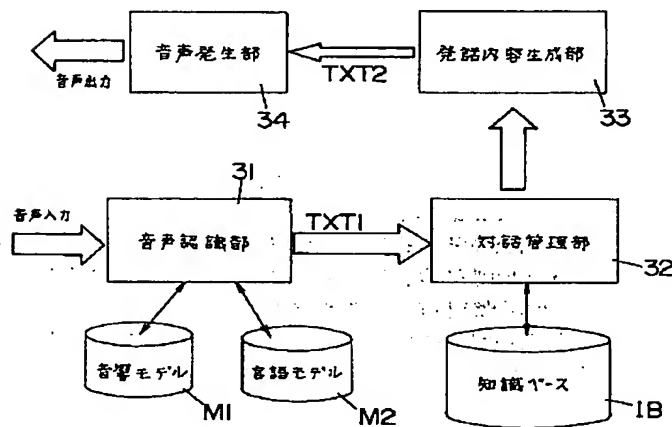
(a)



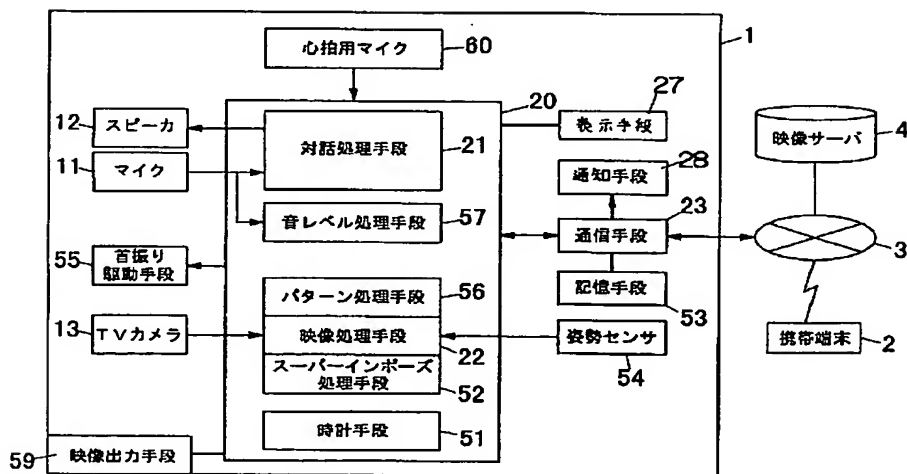
(b)



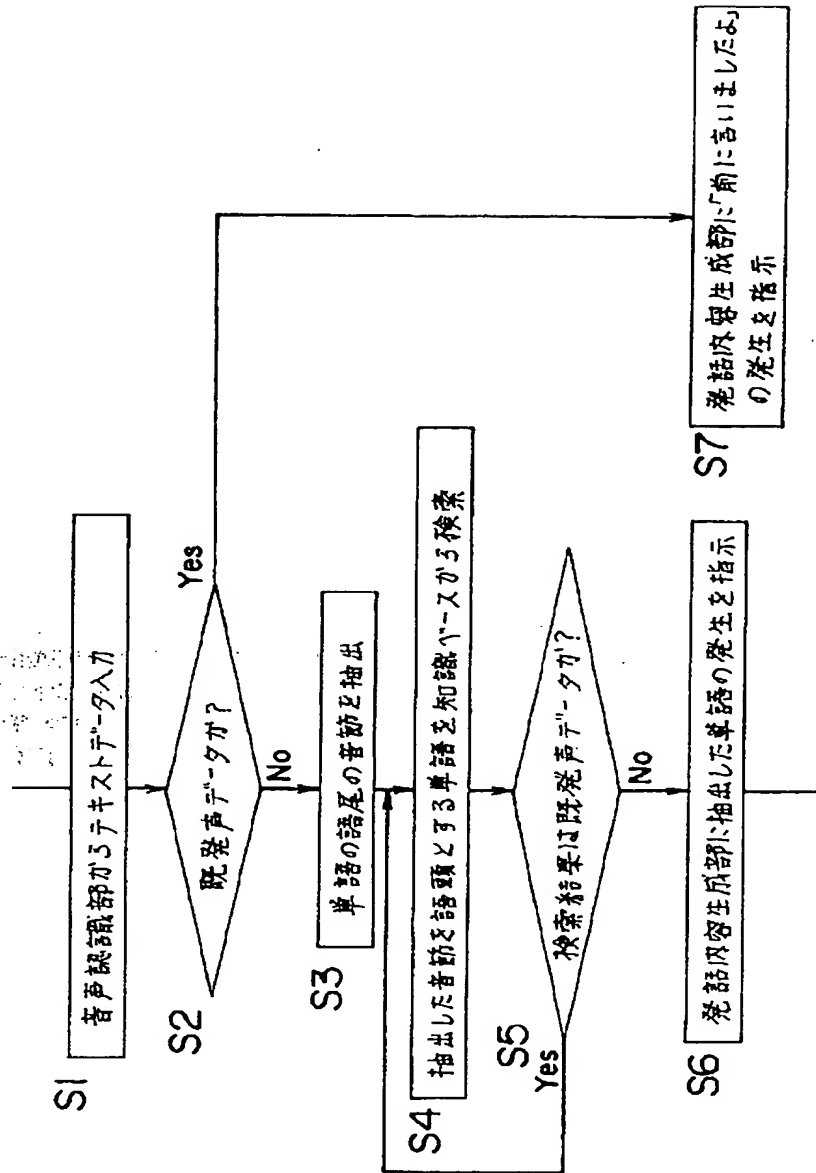
【図2】



【図6】

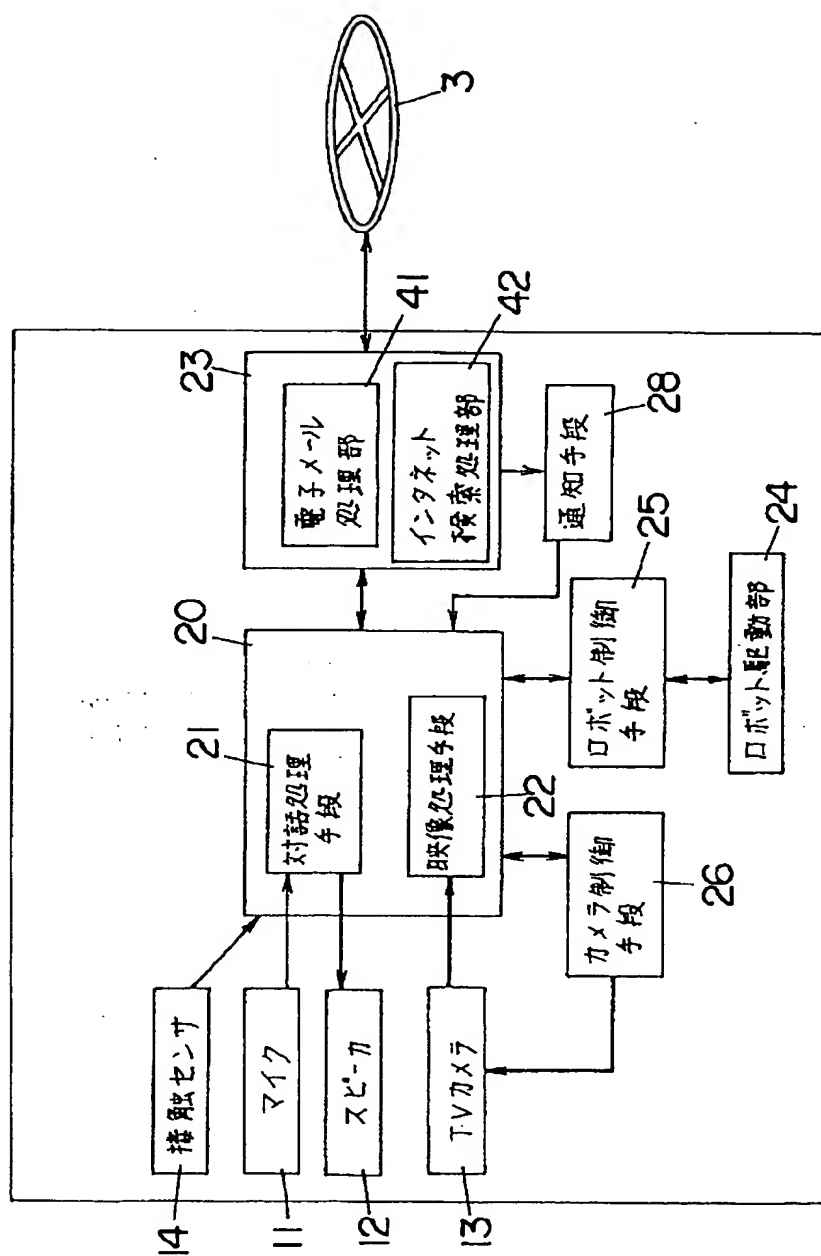


【図3】

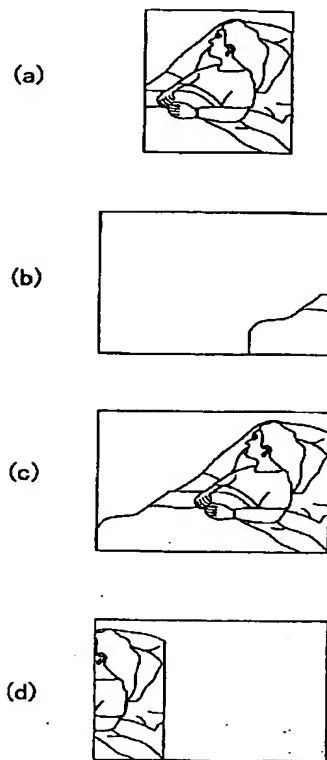




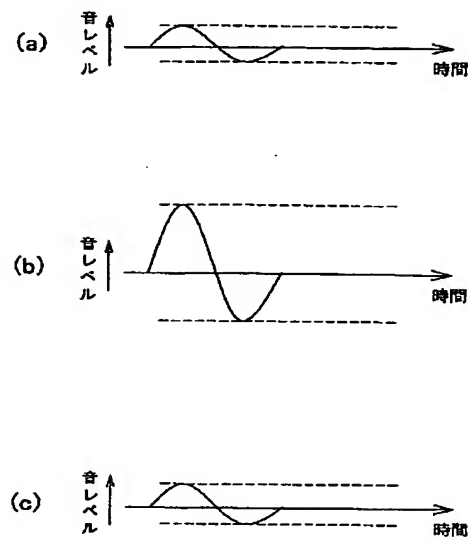
【図4】



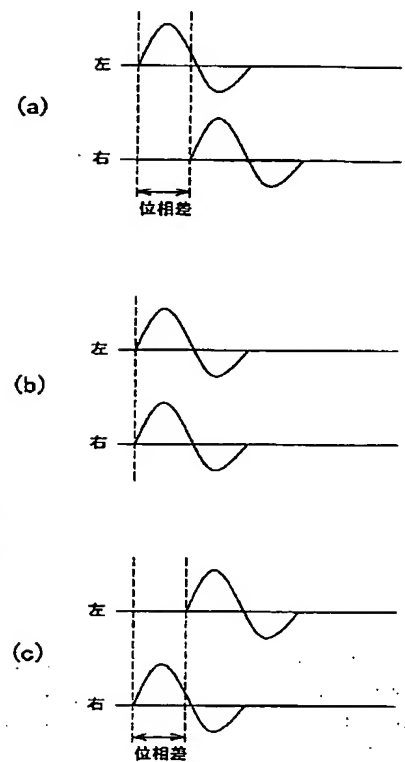
【図8】



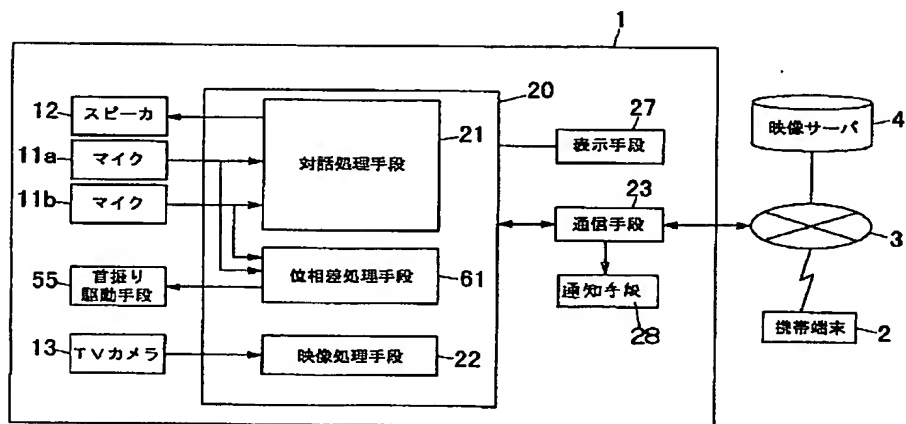
【図9】



【図13】



【図12】



## フロントページの続き

(51)Int. Cl. <sup>7</sup>	識別記号	F I	ターム (参考)
G 1 0 L 15/22		G 1 0 L 3/00	5 5 1 H 5 D 0 4 5
H 0 4 N 7/15			5 7 1 U 5 K 1 0 1
7/18			R

(72)発明者 西山 高史  
大阪府門真市大字門真1048番地松下電工株式会社内

Fターム(参考) 5C054 AA02 CC02 CD03 CE04 CG01  
CH01 CH05 DA01 DA07 EA01  
EA03 EA05 EA07 EH01 FA04  
FC12 FF02 GA04 GB04 GD03  
HA00 HA04  
5C064 AA06 AB04 AC02 AC06 AC12  
AC16 AC22 AD08  
5C086 AA22 BA01 CA09 CA28 CB26  
CB36 DA08 DA14 DA33 EA45  
FA06 FA18  
5C087 AA02 AA08 AA10 AA24 AA25  
AA37 AA44 BB12 BB20 BB46  
BB65 BB74 DD03 DD24 EE05  
FF01 FF02 FF04 FF16 FF19  
FF23 GG02 GG20 GG67  
5D015 KK01 KK02 KK04 LL00  
5D045 AB11  
5K101 KK11 KK13 KK19 LL12 MM07  
NN06 NN08 NN13 NN18 NN21  
RR12